Generation of Plants with Improved Pathogen Resistance and Drought Tolerance

Related Applications

[0001] Priority is claimed under 35 U.S.C. § 120 to international application no. PCT/US03/12981 and U.S. provisional patent application no. 60/482,076, the contents of which are incorporated herein by reference.

Background of the Invention

- [0002] Crop production is affected by numerous environmental factors. Attempts to improve crop yield under stress conditions by plant breeding have been largely unsuccessful, primarily due to the multigenic origin of the adaptive responses (Barkla et al. 1999, Adv Exp Med Biol 464:77-89). Consequently, an increasing amount of research has been dedicated to developing transgenic plants that show increased tolerance to various environmental stresses. For example, transgenic plants have been described having resistance to a broad range of plant pathogens (Stuiver and Custers, 2001, Nature 411:865-8; Melchers and Stuiver, 2000, Curr Opin Plant Biol 3:147-52; Rommens and Kishore, 2000, Curr Opin Biotechnol 11:120-5; and Mourgues et al. 1998, Trends Biotechnol 16:203-10). Transgenic plants displaying increased drought tolerance have also been described (Laporte et al. 2002, J Exp Bot 53:699-705; Qin et al. 2002, Plant Physiol 128:544-51; and Iuchi et al. 2001, Plant J 27:325-33).
- [0003] Activation tagging in plants is a method of generating random mutations by insertion of a heterologous nucleic acid construct comprising regulatory sequences (e.g., an enhancer) into a plant genome. The regulatory sequences can act to enhance transcription of one or more native plant genes. Accordingly, activation tagging is a fruitful method for generating gain-of-function, generally dominant mutants (Hayashi *et al.*, Science (1992) 258: 1350-1353; Weigel *et al.*, Plant Physiology (2000) 122:1003-1013). The inserted construct provides a molecular tag for rapid identification of a native plant gene whose mis-expression causes a mutant phenotype of interest (e.g. drought tolerance or pathogen resistance).

Summary of the Invention

[0004] The invention provides a transgenic plant comprising a plant transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that

- encodes a PRDT1 polypeptide or an ortholog thereof. The transgenic plant is characterized by having increased resistance to pathogens and increased drought tolerance.
- [0005] The present invention further provides a method of producing an altered pathogen resistance and drought tolerance phenotype in a plant. The method comprises introducing into plant progenitor cells a vector comprising a nucleotide sequence that encodes or is complementary to a sequence encoding a PRDT1 polypeptide or an ortholog thereof and growing a transgenic plant that expresses the nucleotide sequence. In one embodiment, the PRDT1 polypeptide has at least 50% sequence identity to the amino acid sequence presented in SEQ ID NO:2 and comprises a SANT domain. In other embodiments, the PRDT1 polypeptide has at least 80% or 90% sequence identity to or has the amino acid sequence presented in SEQ ID NO:2.
- [0006] The invention further provides plants and plant parts obtained by the methods described herein.

Detailed Description of the Invention

Identification of Plants with a PRDT1 Phenotype

- [0007] We used activation tagging in *Arabidopsis* (Weigel *et al*, supra) to generate a large collection of transgenic plants, and tested the plants to identify those that display a phenotype of interest. Briefly, and as further described in the Examples, a large number of *Arabidopsis* plants were mutated with the pSKI015 vector, which comprises a T-DNA from the Ti plasmid of *Agrobacterium tumifaciens*, a viral enhancer element, and a selectable marker gene. When the T-DNA inserts into the genome of transformed plants, the enhancer element can cause up-regulation genes in the vicinity, generally within about 10 kilobase (kb) of the insertion. T1 plants were exposed to the selective agent in order to specifically recover transformed plants that expressed the selectable marker and therefore harbored T-DNA insertions. Seed from transformed plants were planted, and grown under various environmental stress conditions to identify plants having enhanced ability to survive the stress condition.
- [0008] We discovered a transgenic plant line that exhibits increased tolerance to drought conditions. Another plant line was identified that exhibits increased resistance to the plant pathogen, *P. parasitica*. An analysis of the genomic DNA sequence flanking the T-DNA

TAIR designation At1g75250 (GenBank Identifier No. 18410812), which encodes a mybrelated protein, was associated with the drought tolerance phenotype in the one line, and the pathogen resistance phenotype in the other line. We refer to this gene herein as "PRDT1" (for Pathogen Resistant Drought Tolerant). Accordingly, in one aspect of the invention, PRDT1 genes and/or polypeptides may be employed in the development of transgenic plants having an "altered pathogen resistance phenotype", meaning that there is a detectable change in the response of the transgenic plant to pathogenic infection, compared to the similar, but non-modified plant. The phenotype may be apparent in the plant itself (e.g., in growth, viability or particular tissue morphology of the plant) or may be apparent in the ability of the pathogen to proliferate on and/or infect the plant. As used herein, "improved pathogen resistance" refers to increased resistance to a pathogen.

- [0009] In another aspect of the invention the PRDT1 gene/polypeptide may be used to generate transgenic plants having an "altered drought tolerance phenotype", meaning that there is a detectable change in the ability of the transgenic plant to withstand low-water conditions compared to the similar, but non-modified plant. In general, improved (increased) drought tolerance phenotypes (i.e., ability to a plant to survive in low- water conditions that would normally be deleterious to a plant) are of interest.
- [0010] A further aspect of the invention is directed to transgenic plants having a "PRDT1 phenotype" in which the plants exhibit both an increased pathogen resistance phenotype and a drought tolerance phenotype, relative to wild type plants, wherein the plants comprise a plant transformation vector comprising a nucleotide sequence that encodes or is complementary to a sequence that encodes a PRDT1 polypeptide.
- [0011] As used herein, the term "transgenic plant" refers to a plant that comprises a heterologous polynucleotide within its genome. Preferably, the polynucleotide is stably integrated into the genome, such that it is passed on to successive generations. A plant cell, tissue, organ, or any other plant part into which the heterologous polynucleotides have been introduced is considered "transformed", "transfected", or "transgenic". Direct and indirect progeny of transformed plants or plant cells that also contain the heterologous polynucleotide are also considered transgenic. A "heterologous polynucleotide" is a sequence that is not native to the plant cell in which it is expressed. When used to describe a control sequence

(e.g. promoter or enhancer), the term "heterologous" refers to a control sequence (i.e. promoter or enhancer) that does not function in nature to regulate the same gene the expression of which it is currently regulating. Generally, heterologous nucleic acid sequences are not endogenous to the cell or part of the genome in which they are present, and have been added to the cell, by infection, transfection, microinjection, electroporation, or the like. A "heterologous" nucleic acid construct may contain a control sequence/DNA coding sequence combination that is the same as, or different from a control sequence/DNA coding sequence combination found in the native plant.

- [0012] As used herein, the term "plant part" includes any plant organ or tissue, including, without limitation, seeds, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes, pollen, and microspores. Plant cells can be obtained from any plant organ or tissue and cultures prepared therefrom. The class of plants which can be used in the methods of the present invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledenous and dicotyledenous plants.
- [0013] PRDT1 genes may be used in the generation of crops and/or other plant species that have improved ability to survive in low-water conditions, and improved resistance to infection by *P. parasitica* and other oomycetes and may also be useful the generation of plants with improved resistance to fungal, bacterial, and/or other pathogens. Mis-expression of PRDT1 genes may thus reduce the need for fungicides and/or pesticides.

PRDT1 Nucleic Acids and Polypeptides

- [0014] Arabidopsis PRDT1 nucleic acid (coding) sequence is provided in SEQ ID NO:1 and in Genbank entry GI 12331602, nucleotides 20955 21335 (designated F22H5.3 and At1g75250). The corresponding protein sequence is provided in SEQ ID NO:2 and in GI 10092271.
- [0015] As used herein, the term "PRDT1 polypeptide" refers to a full-length PRDT1 protein or a fragment, derivative (variant), or ortholog thereof that is "functionally active," meaning that the protein fragment, derivative, or ortholog exhibits one or more or the functional activities associated with the polypeptide of SEQ ID NO:2. In one preferred embodiment, a functionally active PRDT1 polypeptide causes an altered pathogen resistance and drought

tolerance phenotype when mis-expressed in a plant. In a further preferred embodiment, misexpression of the functionally active PRDT1 polypeptide causes increased resistance to P. parasitica and/or other oomycetes and increased drought tolerance. In another embodiment, a functionally active PRDT1 polypeptide is capable of rescuing defective (including deficient) endogenous PRDT1 activity when expressed in a plant or in plant cells; the rescuing polypeptide may be from the same or from a different species as that with defective activity. In another embodiment, a functionally active fragment of a full length PRDT1 polypeptide (i.e., a native polypeptide having the sequence of SEQ ID NO:2 or a naturally occurring ortholog thereof) retains one of more of the biological properties associated with the full-length PRDT1 polypeptide, such as signaling activity, binding activity, catalytic activity, or cellular or extra-cellular localizing activity. Some preferred PRDT1 polypeptides display DNA binding activity. A PRDT1 fragment preferably comprises a PRDT1 domain, such as a C- or N-terminal or catalytic domain, among others, and preferably comprises at least 10, preferably at least 20, more preferably at least 25, and most preferably at least 50 contiguous amino acids of a PRDT1 protein. Functional domains can be identified using the PFAM program (Bateman A et al., 1999 Nucleic Acids Res 27:260-262; website at pfam.wustl.edu). A preferred PRDT1 fragment comprises a SANT domain (SM00395) identified by PFAM, at approximately amino acids 8-60. Functionally active variants of fulllength PRDT1 polypeptides or fragments thereof include polypeptides with amino acid insertions, deletions, or substitutions that retain one of more of the biological properties associated with the full-length PRDT1 polypeptide. In some cases, variants are generated that change the post-translational processing of a PRDT1 polypeptide. For instance, variants may have altered protein transport or protein localization characteristics or altered protein half-life compared to the native polypeptide.

- [0016] As used herein, the term "PRDT1 nucleic acid" encompasses nucleic acids with the sequence provided in or complementary to the sequence provided in SEQ ID NO:1, as well as functionally active fragments, derivatives, or orthologs thereof. A PRDT1 nucleic acid of this invention may be DNA, derived from genomic DNA or cDNA, or RNA.
- [0017] In one embodiment, a functionally active PRDT1 nucleic acid encodes or is complementary to a nucleic acid that encodes a functionally active PRDT1 polypeptide.
 Included within this definition is genomic DNA that serves as a template for a primary RNA

transcript (i.e., an mRNA precursor) that requires processing, such as splicing, before encoding the functionally active PRDT1 polypeptide. A PRDT1 nucleic acid can include other non-coding sequences, which may or may not be transcribed; such sequences include 5' and 3' UTRs, polyadenylation signals and regulatory sequences that control gene expression, among others, as are known in the art. Some polypeptides require processing events, such as proteolytic cleavage, covalent modification, etc., in order to become fully active.

Accordingly, functionally active nucleic acids may encode the mature or the pre-processed PRDT1 polypeptide, or an intermediate form. A PRDT1 polynucleotide can also include heterologous coding sequences, for example, sequences that encode a marker included to facilitate the purification of the fused polypeptide, or a transformation marker.

- [0018] In another embodiment, a functionally active PRDT1 nucleic acid is capable of being used in the generation of loss-of-function pathogen resistance or drought tolerance phenotypes, for instance, via antisense suppression, co-suppression, etc.
- [0019] In one preferred embodiment, a PRDT1 nucleic acid used in the methods of this invention comprises a nucleic acid sequence that encodes or is complementary to a sequence that encodes a PRDT1 polypeptide having at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95% or more sequence identity to the polypeptide sequence presented in SEQ ID NO:2.
- [0020] In another embodiment a PRDT1 polypeptide of the invention comprises a polypeptide sequence with at least 50% or 60% identity to the PRDT1 polypeptide sequence of SEQ ID NO:2, and may have at least 70%, 80%, 85%, 90% or 95% or more sequence identity to the PRDT1 polypeptide sequence of SEQ ID NO:2. In another embodiment, a PRDT1 polypeptide comprises a polypeptide sequence with at least 50%, 60%, 70%, 80%, 85%, 90% or 95% or more sequence identity to a functionally active fragment of the polypeptide presented in SEQ ID NO:2, such as a SANT domain. In yet another embodiment, a PRDT1 polypeptide comprises a polypeptide sequence with at least 50%, 60 %, 70%, 80%, or 90% identity to the polypeptide sequence of SEQ ID NO:2 over its entire length and comprises a SANT domain.
- [0021] In another aspect, a PRDT1 polynucleotide sequence is at least 50% to 60% identical over its entire length to the PRDT1 nucleic acid sequence presented as SEQ ID NO:1, or nucleic acid sequences that are complementary to such a PRDT1 sequence, and may comprise at least 70%, 80%, 85%, 90% or 95% or more sequence identity to the PRDT1

sequence presented as SEQ ID NO:1 or a functionally active fragment thereof, or complementary sequences.

As used herein, "percent (%) sequence identity" with respect to a specified subject [0022] sequence, or a specified portion thereof, is defined as the percentage of nucleotides or amino acids in the candidate derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul et al., J. Mol. Biol. (1990) 215:403-410; website at blast.wustl.edu/blast/README.html) with search parameters set to default values. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A "% identity value" is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation. A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the protein is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

[0023] Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that hybridize to the nucleic acid sequence of SEQ ID NO:1. The stringency of hybridization can be controlled by temperature, ionic strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing. Conditions routinely used are well known (see, e.g., Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook et al., supra). In some

embodiments, a nucleic acid molecule of the invention is capable of hybridizing to a nucleic acid molecule containing the nucleotide sequence of SEQ ID NO:1 under stringent hybridization conditions that comprise: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X Denhardt's solution, 0.05% sodium pyrophosphate and 100 μg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C in a solution containing 6X SSC, 1X Denhardt's solution, 100 μg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C for 1 h in a solution containing 0.2X SSC and 0.1% SDS (sodium dodecyl sulfate). In other embodiments, moderately stringent hybridization conditions are used that comprise: pretreatment of filters containing nucleic acid for 6 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA; hybridization for 18-20 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 μg/ml salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C in a solution containing 2X SSC and 0.1% SDS. Alternatively, low stringency conditions can be used that comprise: incubation for 8 hours to overnight at 37° C in a solution comprising 20% formamide, 5 x SSC, 50 mM sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μ g/ml denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1 x SSC at about 37° C for 1 hour.

- [0024] As a result of the degeneracy of the genetic code, a number of polynucleotide sequences encoding a PRDT1 polypeptide can be produced. For example, codons may be selected to increase the rate at which expression of the polypeptide occurs in a particular host species, in accordance with the optimum codon usage dictated by the particular host organism (see, e.g., Nakamura Y et al, Nucleic Acids Res (1999) 27:292). Such sequence variants may be used in the methods of this invention.
- [0025] The methods of the invention may use orthologs of the *Arabidopsis* PRDT1.

 Methods of identifying the orthologs in other plant species are known in the art. Normally, orthologs in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. In evolution, when a gene duplication event follows

speciation, a single gene in one species, such as Arabidopsis, may correspond to multiple genes (paralogs) in another. As used herein, the term "orthologs" encompasses paralogs. When sequence data is available for a particular plant species, orthologs are generally identified by sequence homology analysis, such as BLAST analysis, usually using protein bait sequences. Sequences are assigned as a potential ortholog if the best hit sequence from the forward BLAST result retrieves the original query sequence in the reverse BLAST (Huynen MA and Bork P, Proc Natl Acad Sci (1998) 95:5849-5856; Huynen MA et al., Genome Research (2000) 10:1204-1210). Programs for multiple sequence alignment, such as CLUSTAL (Thompson JD et al, 1994, Nucleic Acids Res 22:4673-4680) may be used to highlight conserved regions and/or residues of orthologous proteins and to generate phylogenetic trees. In a phylogenetic tree representing multiple homologous sequences from diverse species (e.g., retrieved through BLAST analysis), orthologous sequences from two species generally appear closest on the tree with respect to all other sequences from these two species. Structural threading or other analysis of protein folding (e.g., using software by ProCeryon, Biosciences, Salzburg, Austria) may also identify potential orthologs. Nucleic acid hybridization methods may also be used to find orthologous genes and are preferred when sequence data are not available. Degenerate PCR and screening of cDNA or genomic DNA libraries are common methods for finding related gene sequences and are well known in the art (see, e.g., Sambrook, supra; Dieffenbach and Dveksler (Eds.) PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY, 1989). For instance, methods for generating a cDNA library from the plant species of interest and probing the library with partially homologous gene probes are described in Sambrook et al. A highly conserved portion of the Arabidopsis PRDT1 coding sequence may be used as a probe. PRDT1 ortholog nucleic acids may hybridize to the nucleic acid of SEQ ID NO:1 under high, moderate, or low stringency conditions. After amplification or isolation of a segment of a putative ortholog, that segment may be cloned and sequenced by standard techniques and utilized as a probe to isolate a complete cDNA or genomic clone. Alternatively, it is possible to initiate an EST project to generate a database of sequence information for the plant species of interest. In another approach, antibodies that specifically bind known PRDT1 polypeptides are used for ortholog isolation. Western blot analysis can determine that a PRDT1 ortholog (i.e., an orthologous protein) is present in a crude extract of a particular

plant species. When reactivity is observed, the sequence encoding the candidate ortholog may be isolated by screening expression libraries representing the particular plant species. Expression libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Sambrook, et al., supra. Once the candidate ortholog(s) are identified by any of these means, candidate orthologous sequence are used as bait (the "query") for the reverse BLAST against sequences from Arabidopsis or other species in which PRDT1 nucleic acid and/or polypeptide sequences have been identified.

- [0026] PRDT1 nucleic acids and polypeptides may be obtained using any available method. For instance, techniques for isolating cDNA or genomic DNA sequences of interest by screening DNA libraries or by using polymerase chain reaction (PCR), as previously described, are well known in the art. Alternatively, nucleic acid sequence may be synthesized. Any known method, such as site directed mutagenesis (Kunkel TA *et al.*, Methods Enzymol. (1991) 204:125-39), may be used to introduce desired changes into a cloned nucleic acid.
- [0027] In general, the methods of the invention involve incorporating the desired form of the PRDT1 nucleic acid into a plant expression vector for transformation of in plant cells, and the PRDT1 polypeptide is expressed in the host plant.
- [0028] An isolated PRDT1 nucleic acid molecule is other than in the form or setting in which it is found in nature and is identified and separated from least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRDT1 nucleic acid. However, an isolated PRDT1 nucleic acid molecule includes PRDT1 nucleic acid molecules contained in cells that ordinarily express PRDT1 where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

Generation of Transgenic Plants with a PRDT1 Phenotype

[0029] PRDT1 nucleic acids and polypeptides may be used in the generation of transgenic plants having a modified pathogen resistance phenotype; in general, improved resistance phenotypes are of interest. Pathogenic infection may affect seeds, fruits, blossoms, foliage, stems, tubers, roots, etc. Accordingly, resistance may be observed in any part of the plant. In a preferred embodiment, altered expression of the PRDT1 gene in a plant is used to generate plants with increased resistance to *P. parasitica*. In a further preferred embodiment,

plants that mis-express PRDT1 may also display altered resistance to other pathogens. Other oomycete pathogens of interest include Pythium spp, Phytophthora spp, Bremia lactucae, Peronosclerospora spp., Pseudoperonospora. Sclerophthora macrospora, Sclerospora graminicola, Plasmopara viticola, and Albugo candidia. Fungal pathogens of interest include Alternaria brassicicola, Botrytis cinerea, Erysiphe cichoracearum, Fusarium oxysporum, Plasmodiophora brassica, Rhizoctonia solani, Colletotrichum coccode, Sclerotinia spp., Aspergillus spp., Penicillium spp., Ustilago spp., and Tilletia spp. Bacterial pathogens of interest include Agrobacterium tumefaciens, Erwinia tracheiphila, Erwinia stewartii, Xanthomonas phaseoli, Erwinia amylovora, Erwinia carotovora, Pseudomonas syringae, Pelargonium spp, Pseudomonas cichorii, Xanthomonas fragariae, Pseudomonas morsprunorum, Xanthomonas campestris.

- [0030] PRDT1 nucleic acids and polypeptides may also be used in the generation of transgenic plants having a modified, preferably an improved drought tolerance phenotype. Such plants may further display increased tolerance to other abiotic stresses, particular salt-stress and freezing, as responses to these stresses and drought stress are mediated by ABA (Thomashow, 1999 Annu. Revl Plant Physiol. Plant Mol. Biol 50: 571; Cushman and Bohnert, 2000, Curr. Opin. Plant Biol. 3: 117; Kang et al. 2002, Plant Cell 14:343-357; Quesada et al. 2000, Genetics 154: 421; Kasuga et al. 1999, Nature Biotech. 17: 287-291).
- [0031] The methods described herein are generally applicable to all plants. Although activation tagging and gene identification is carried out in *Arabidopsis*, the PRDT1 gene (or an ortholog, variant or fragment thereof) may be expressed in any type of plant. In preferred embodiments, the invention is directed to crops including maize, soybean, cotton, rice, wheat, barley, tomato, canola, turfgrass, sugar beets, onions, beans, and flax. Other crops include alfalfa, tobacco, and other forage crops. The invention may also be directed to fruit-and vegetable-bearing plants, plants used in the cut flower industry, grain-producing plants, oil-producing plants, and nut-producing plants, among others.
- [0032] The skilled artisan will recognize that a wide variety of transformation techniques exist in the art, and new techniques are continually becoming available. Any technique that is suitable for the target host plant can be employed within the scope of the present invention. For example, the constructs can be introduced in a variety of forms including, but not limited to as a strand of DNA, in a plasmid, or in an artificial chromosome. The introduction of the

constructs into the target plant cells can be accomplished by a variety of techniques, including, but not limited to *Agrobacterium*-mediated transformation, electroporation, microinjection, microprojectile bombardment calcium-phosphate-DNA co-precipitation or liposome-mediated transformation of a heterologous nucleic acid. The transformation of the plant is preferably permanent, *i.e.* by integration of the introduced expression constructs into the host plant genome, so that the introduced constructs are passed onto successive plant generations. Depending upon the intended use, a heterologous nucleic acid construct comprising a PRDT1 polynucleotide may encode the entire protein or a biologically active portion thereof.

- [0033] In one embodiment, binary Ti-based vector systems may be used to transfer polynucleotides. Standard *Agrobacterium* binary vectors are known to those of skill in the art, and many are commercially available (e.g., pBI121 Clontech Laboratories, Palo Alto, CA).
- [0034] The optimal procedure for transformation of plants with Agrobacterium vectors will vary with the type of plant being transformed. Exemplary methods for Agrobacterium-mediated transformation include transformation of explants of hypocotyl, shoot tip, stem or leaf tissue, derived from sterile seedlings and/or plantlets. Such transformed plants may be reproduced sexually, or by cell or tissue culture. Agrobacterium transformation has been previously described for a large number of different types of plants and methods for such transformation may be found in the scientific literature.
- [0035] Expression (including transcription and translation) of PRDT1 may be regulated with respect to the level of expression, the tissue type(s) where expression takes place and/or developmental stage of expression. A number of heterologous regulatory sequences (e.g., promoters and enhancers) are available for controlling the expression of a PRDT1 nucleic acid. These include constitutive, inducible and regulatable promoters, as well as promoters and enhancers that control expression in a tissue- or temporal-specific manner. Exemplary constitutive promoters include the raspberry E4 promoter (U.S. Patent Nos. 5,783,393 and 5,783,394), the 35S CaMV (Jones JD et al, Transgenic Res (1992) 1:285-297), the CsVMV promoter (Verdaguer B et al., Plant Mol Biol (1998) 37:1055-1067) and the melon actin promoter (published PCT application WO0056863). Exemplary tissue-specific promoters include the tomato E4 and E8 promoters (U.S. Patent No. 5,859,330) and the tomato 2AII

gene promoter (Van Haaren MJJ et al., Plant Mol Bio (1993) 21:625-640). In one preferred embodiment, PRDT1 expression is under the control of a pathogen-inducible promoter (Rushton et al., The Plant Cell (2002) 14:749-762). In another preferred embodiment, PRDT1 expression is under control of regulatory sequences from genes whose expression is associated with drought stress. For example, when the promoter of the drought stress responsive Arabidopsis rd29A gene was used to drive expression of DREB1A, Arabidopsis plants were more tolerant to drought, salt and freezing stress and did not have the stunted stature associated with plants over-expressing the DREB1A gene from the CaMV 35S promoter (Kasuga et al, 1999 Nature Biotech 17: 287). Promoters from other Arabidopsis genes that are responsive to drought stress, such as COR47 (Welin et al. 1995, Plant Mol. Biol. 29: 391), KIN1 (Kurkela and Franck, 1990, Plant Mol. Biol. 15: 137), RD22BP (Abe et al. 1997, Plant Cell 9, 1859), ABAI (Accession Number AAG17703), and ABA3 (Xiong et al. 2001, Plant Cell 13: 2063), could be used. Promoters from drought stress inducible genes in other species could be used also. Examples are the rab17, ZmFer1 and ZmFer2 genes from maize (Bush et al, 1997 Plant J 11:1285; Fobis-Loisy, 1995 Eur J Biochem 231:609), the tdi-65 gene from tomato (Harrak, 2001 Genome 44:368), the His1 gene of tobacco (Wei and O'Connell, 1996 Plant Mol Biol 30:255), the Vupat1 gene from cowpea (Matos, 2001 FEBS Lett 491:188), and CDSP34 from Solanum tuberosum (Gillet et al, 1998 Plant J 16:257).

[0036] In yet another aspect, in some cases it may be desirable to inhibit the expression of endogenous PRDT1 in a host cell. Exemplary methods for practicing this aspect of the invention include, but are not limited to antisense suppression (Smith, et al., Nature (1988) 334:724-726; van der Krol et al., Biotechniques (1988) 6:958-976); co-suppression (Napoli, et al., Plant Cell (1990) 2:279-289); ribozymes (PCT Publication WO 97/10328); and combinations of sense and antisense (Waterhouse, et al., Proc. Natl. Acad. Sci. USA (1998) 95:13959-13964). Methods for the suppression of endogenous sequences in a host cell typically employ the transcription or transcription and translation of at least a portion of the sequence to be suppressed. Such sequences may be homologous to coding as well as noncoding regions of the endogenous sequence. Antisense inhibition may use the entire cDNA sequence (Sheehy et al., Proc. Natl. Acad. Sci. USA (1988) 85:8805-8809), a partial cDNA sequence including fragments of 5' coding sequence, (Cannon et al., Plant Molec. Biol. (1990) 15:39-47), or 3' non-coding sequences (Ch'ng et al., Proc. Natl. Acad. Sci. USA

- (1989) 86:10006-10010). Cosuppression techniques may use the entire cDNA sequence (Napoli et al., *supra*; van der Krol et al., The Plant Cell (1990) 2:291-299), or a partial cDNA sequence (Smith et al., Mol. Gen. Genetics (1990) 224:477-481).
- [0037] Standard molecular and genetic tests may be performed to further analyze the association between a gene and an observed phenotype. Exemplary techniques are described below.

1. DNA/RNA analysis

- [0038] The stage- and tissue-specific gene expression patterns in mutant versus wild-type lines may be determined, for instance, by in situ hybridization. Analysis of the methylation status of the gene, especially flanking regulatory regions, may be performed. Other suitable techniques include overexpression, ectopic expression, expression in other plant species and gene knock-out (reverse genetics, targeted knock-out, viral induced gene silencing [VIGS, see Baulcombe D, (1999) Arch Virol Suppl 15:189-201]).
- [0039] In a preferred application expression profiling, generally by microarray analysis, is used to simultaneously measure differences or induced changes in the expression of many different genes. Techniques for microarray analysis are well known in the art (Schena M et al., Science (1995) 270:467-470; Baldwin D et al., Cur Opin Plant Biol. (1999) 2(2):96-103; Dangond F, Physiol Genomics (2000) 2:53-58; van Hal NL et al., J Biotechnol (2000) 78:271-280; Richmond T and Somerville S, Curr Opin Plant Biol (2000) 3:108-116). Expression profiling of individual tagged lines may be performed. Such analysis can identify other genes that are coordinately regulated as a consequence of the overexpression of the gene of interest, which may help to place an unknown gene in a particular pathway.

2. Gene Product Analysis

[0040] Analysis of gene products may include recombinant protein expression, antisera production, immunolocalization, biochemical assays for catalytic or other activity, analysis of phosphorylation status, and analysis of interaction with other proteins via yeast two-hybrid assays.

3. Pathway Analysis

[0041] Pathway analysis may include placing a gene or gene product within a particular biochemical, metabolic or signaling pathway based on its mis-expression phenotype or by sequence homology with related genes. Alternatively, analysis may comprise genetic crosses

with wild-type lines and other mutant lines (creating double mutants) to order the gene in a pathway, or determining the effect of a mutation on expression of downstream "reporter" genes in a pathway.

Generation of Mutated Plants with a PRDT1 Phenotype

- [0042] The invention further provides a method of identifying plants that have mutations in endogenous PRDT1 that confer increased pathogen resistance and drought tolerance, and generating pathogen-resistant and drought-tolerant progeny of these plants that are not genetically modified. In one method, called "TILLING" (for targeting induced local lesions in genomes), mutations are induced in the seed of a plant of interest, for example, using EMS treatment. The resulting plants are grown and self-fertilized, and the progeny are used to prepare DNA samples. PRDT1-specific PCR are used to identify whether a mutated plant has a PRDT1 mutation. Plants having PRDT1 mutations may then be tested for the PRDT1 phenotype, or alternatively, plants may be tested for the PRDT1 phenotype, and then PRDT1-specific PCR is used to determine whether a plant having increased pathogen resistance and drought tolerance has a mutated PRDT1 gene. TILLING can identify mutations that may alter the expression of specific genes or the activity of proteins encoded by these genes (see Colbert et al (2001) Plant Physiol 126:480-484; McCallum et al (2000) Nature Biotechnology 18:455-457).
- [0043] In another method, a candidate gene/Quantitative Trait Locus (QTLs) approach can be used in a marker-assisted breeding program to identify alleles of or mutations in the PRDT1 gene or orthologs of PRDT1 that may confer the PRDT1 phenotype (see Foolad et al., Theor Appl Genet. (2002) 104(6-7):945-958; Rothan et al., Theor Appl Genet (2002) 105(1):145-159); Dekkers and Hospital, Nat Rev Genet. (2002) Jan;3(1):22-32). Thus, in a further aspect of the invention, a PRDT1 nucleic acid is used to identify whether a plant having increased pathogen resistance and/or drought tolerance has a mutation in endogenous PRDT1 or has a particular allele that causes the increased pathogen resistance and/or drought tolerance.
- [0044] All publications cited herein are expressly incorporated herein by reference for the purpose of describing and disclosing compositions and methodologies that might be used in

connection with the invention. All cited patents, patent applications, and sequence information in referenced websites and public databases are also incorporated by reference.

Examples

EXAMPLE 1

Activation Tagging in Arabidopsis

[0045] Mutants were generated using the activation tagging "ACTTAG" vector, pSKI015 (GI 6537289; Weigel D et al., 2000). Standard methods were used for the generation of Arabidopsis transgenic plants, and were essentially as described in published application PCT WO0183697. Briefly, To Arabidopsis (Col-0) plants were transformed with Agrobacterium carrying the pSKI015 vector, which comprises T-DNA derived from the Agrobacterium Ti plasmid, an herbicide resistance selectable marker gene, and the 4X CaMV 35S enhancer element. Transgenic plants were selected at the T1 generation based on herbicide resistance. T2 seed was collected from T1 plants and stored in an indexed collection, and a portion of the T2 seed was accessed for the screen.

EXAMPLE 2

Pathogen Resistance Screen

- [0046] Approximately 18 T2 seeds from each of the greater than 40,00 lines generated in Example 1, were planted in soil. The seed were stratified for three days and then grown in the greenhouse for seven days. The seedlings were inoculated with approximately 1x10⁵ conidia per ml *P. parasitica* spores and incubated in a dew room at 18°C and 100% humidity for 24 hours. The plants were then moved to a growth room at 20°C and 60% relative humidity with ten-hour long light period for six days. Individual plants were evaluated for the presence or absence of conidiophores on cotyledons. Lines in which at least a single plant showed no conidiophore growth were re-tested in a secondary screen by releasing three sets of 18 seed and screening for resistance to *P. parasitica* growth as before.
- [0047] Lines in which a significant number of plants showed no conidiophores after infection were subjected to a tertiary screen. Approximately 54 T2 seed were released, planted individually and infected with *P. parasitica* as before. The plants were evaluated for the number of conidiophores growing on a single cotyledon and ranked by the following scoring

- system: a score of 0 indicates 0 conidiophores per cotyledon, 1 indicates 1-5 conidiophores per cotyledon, 2 indicates 6-10 conidiophores per cotyledon, 3 indicates 11-20 conidiophores per cotyledon, and 4 indicates greater than 20 conidiophores per cotyledon
- [0048] The ACTTAG line designated W000058335 was identified as having an increased resistance phenotype. Specifically, 15.2% of individual plants showed no conidiophores in the secondary screen. In the tertiary screen, 31 plants scored as 0 (39.2%), 31 as 1 (39.2%), 4 as 2 (5.1%), 10 as 3 (12.7%) and 3 as 4 (3.8%). Control wild-type Col-0 plants were more susceptible; 36 plants scored 0 (7.6%), 21 as 1 (4.4%), 79 as 2 (16.6%), 250 as 3 (52.5%) and 90 as 4 (18.9%).
- [0049] We performed standard molecular analyses, essentially as described in patent application PCT WO0183697, to determine the site of the T-DNA insertion associated with the increased pathogen resistance phenotype. Briefly, genomic DNA was extracted from plants exhibiting increased pathogen resistance. PCR, using primers specific to the pSKI015 vector, confirmed the presence of the 35S enhancer in plants from line W000058335, and Southern blot analysis verified the genomic integration of the ACTTAG T-DNA and showed the presence of a single T-DNA insertion in the transgenic line.
- [0050] Plasmid rescue and inverse PCR were used to recover genomic DNA flanking the T-DNA insertion, which was then subjected to sequence analysis.
- [0051] The sequence flanking the right T-DNA border was subjected to a basic BLASTN search and/or a search of the *Arabidopsis* Information Resource (TAIR) database (available at the arabidopsis.org website), which revealed sequence identity to BAC F22H5, (GI 12331602), mapped to chromosome 1. The junction of the left border of the T-DNA is at nt 20167 of F22H5, and the right border junction is at nt 20229. Sequence analysis revealed that the T-DNA had inserted in the vicinity (*i.e.*, within about 10 kb) of the gene whose nucleotide sequence is presented as SEQ ID NO: 1 (GI 18410812 and GI 12331602, nucleotides 20955 21335) and which we designated PRDT1. Specifically, the right border was approximately 500 bp upstream of the start codon of SEQ ID NO:1.

Drought Tolerance Screen

- [0052] Drought stress was imposed on transgenic plants generated in Example 1 by withholding water for 21-25 days. At this time the plants were beginning to bolt. From 6 to 10 days after the initiation of drought stress, observations of the plants were taken daily. Transgenic plants were compared to each other and wild-type control plants. Putative drought tolerant lines were identified as containing at least 2 plants that remained green and viable after wild-type plants had died or as plants that contained a soil moisture content of at least 50 mV (using a Delt-T Devices HH2 Soil Moisture Meter with a ML2x Theta Probe).
- [0053] After drought tolerant lines were identified, water was applied to the plants to allow them to recover. When possible, T3 seed was collected from the plants. This T3 seed was then grown and the plants assessed for drought tolerant phenotype as described above.
- [0054] The drought tolerant plants were then subjected to an "excised leaf transpiration test" in which seeds were planted, stratified, and grown for three weeks as described above. Then, either the entire rosette or a single leaf was excised and placed on a pre-weighed plastic weigh dish and left on the bench at room temperature. The mass of the plant material was recorded immediately after excision and at 30 min intervals afterward. The mass of drought tolerant plants often decreased less rapidly indicating that they were transpiring less rapidly.
- [0055] To detect lines containing or lacking the insert, PCR analysis was performed using a set of DNA oligonucleotide primers; one that hybridizes to sequences in pSKI015, the other that hybridizes to sequences flanking the insert. Genotyping of individuals analyzed drought tolerance experiments indicated that plants containing the insert identified in plant line W000114956 were more tolerant of drought stress than plants without the insert.
- [0056] The same methods described in Example 2 were used to determine the site of the T-DNA insertion associated with the increased pathogen resistance phenotype. Sequence analysis revealed that the T-DNA had inserted 423 base pairs from the start codon in At1g75250, which is the same gene that was responsible for the pathogen resistance phenotype of plant line W000058335, as described in Example 2.

Analysis of Arabidopsis PRDT1 Sequence

- [0057] The amino acid sequence predicted from the PRDT1 nucleic acid sequence is presented in SEQ ID NO:2 and GI 10092271. Sequence analyses were performed with BLAST (Altschul *et al.*, 1997, J. Mol. Biol. 215:403-410), PFAM (Bateman *et al.*, 1999), PSORT (Nakai K, and Horton P, 1999, Trends Biochem Sci 24:34-6), and CLUSTALW (Thompson JD et al, 1994, Nucleic Acids Res 22:4673-4680), among others. The PRDT1 protein has been characterized as a myb-related protein. PFAM analysis indicated a SANT DNA-binding domain at approximately amino acids 8-60.
- [0058] The retroviral oncogene v-myb, and its cellular counterpart c-myb, encode nuclear DNA-binding proteins (Klempnauer and Sippel, 1987, EMBO J. 6: 2719-2725; Biednkapp et al. 1988, Nature 335: 835-837). These belong to the SANT domain family that specifically recognize the sequence YAAC(G/T)G (Aasland et al. 1996, Trends Biochem. Sci. 21:87-88). In myb, one of the most conserved regions consisting of three tandem repeats has been shown to be involved in DNA-binding.
- [0059] Analysis using BLASTP or TBLASTN identified a number of related proteins and proteins predicted from nucleic acid (generally EST) sequences in other plant species.

 Related sequences, which are candidate orthologs, are presented in SEQ ID NOs 3-14 and descriptions from GenBank are provided below:
- [0060] SEQ ID NO:3 translation, gi|887283|gb|L38243.1|L38243 BNAF0581E Mustard flower buds Brassica rapa cDNA ORF 98aa Brassica rapa
- [0061] SEQ ID NO:4 translation, gi|18459015|gb|BM437293.1|BM437293 VVA017C08_54081 An expressed sequence tag database for abiotic st 75aa Vitis vinifera
- [0062] SEQ ID NO:5 translation, gi|15288211|gb|BI472102.1|BI472102 sah99e03.y1 Gm-c1050 Glycine max cDNA clone GENOME SYSTEMS CLONE 97aa Glycine max
- [0063] SEQ ID NO:6 translation, gi|15258392|gb|BI433702.1|BI433702 EST536463 P. infestans-challenged leaf Solanum tuberosum cDNA clo 88aa Solanum tuberosum
- [0064] SEQ ID NO: translation, gi|14492357|gb|BI071737.1|BI071737 C063P09U Populus strain T89 leaves Populus tremula x Populus trem 71aa P o
- [0065] SEQ ID NO:8 translation, gi|7981380|emb|CAB91874.1| (AJ277944) myb-related protein [Lycopersicon esculentum] 88aa Lycopersicon esculentum

- [0066] SEQ ID NO:9 gi|5091605|gb|AAD39594.1|AC007858_8 (AC007858) 10A19I.9 [Oryza sativa] 126aa Oryza sativa
- [0067] SEQ ID NO:10 gi|5091604|gb|AAD39593.1|AC007858_7 (AC007858) 10A19I.8 [Oryza sativa] 236aa Oryza sativa
- [0068] SEQ ID NO:11 gi|18394750|ref|NP_564087.1| (NM_101808) myb-related protein, putative [Arabidopsis thaliana] 92aa Arabidopsis thaliana
- [0069] SEQ ID NO:12 gi|15226604|ref|NP_179759.1| (NM_127736) unknown protein [Arabidopsis thaliana] [gi|4567225|gb|AAD236 101aa Arabidopsis thaliana]
- [0070] SEQ ID NO:13 gi|15234999|ref|NP_195636.1| (NM_120086) putative protein [Arabidopsis thaliana] gi|7487341|pir||T08 97aa Arabidopsis thaliana
- [0071] SEQ ID NO:14 gi|8778436|gb|AAF79444.1|AC025808_26 F18O14.26 [Arabidopsis thaliana]
- [0072] SEQ ID NO:15 gi|20161824|dbj|BAB90739.1|[20161824] P0663E10.22 [Oryza sativa (japonica cultivar-group)]
- [0073] SEQ ID NO:16 gi|19387259|gb|AAL87171.1|AF480496_25[19387259] putative myb-related protein [Oryza sativa (japonica cultivar-group)]
- [0074] The predicted protein from an EST contig identified from *Lycopersicon esculentum* was also identified as an ortholog having 58% sequence identity to SEQ ID NO:2, and is presented as SEQ ID NO:17.
- [0075] The translation of an EST contig from *Solanum tuberosum*, presented as SEQ ID NO:18, was identified as having 53% sequence identify to SEQ ID NO:1.

Confirmation of Pathogen Resistance Phenotype/Genotype Association

[0076] PCR analysis, using primers to sequences in pSKI015 or flanking the insert, was used to detect lines containing or lacking the insert. W000058335 individuals analyzed in the tertiary screen were genotyped. Results indicated that plants that were homozygous or hemizygous for the insert were more resistant to *P. parasitica* infection than plants that were homozygous wild-type; 100% of the plants homozygous for the insert and 97% of the plants hemizygous for the insertion received resistance scores of 0 or 1 while only 31% of the wild-

type segregants scored 0 or 1. These results suggest that the *P. parasitica* resistance trait in W000058335 is caused by the overexpression of *PRDT1* and is inherited in a dominant manner

[0077] RT-PCR analysis showed that the PRDT1 gene was overexpressed in plants from the line displaying the *P. parasitica* resistance phenotype. Specifically, RNA was extracted from tissues derived from plants exhibiting the resistance phenotype and from wild type COL-0 plants. RT-PCR was performed using primers specific to the sequence presented as SEQ ID NO:1, to other predicted genes in the vicinity of the T-DNA insertion (At1g75240, At1g75260, and At1g75270), and to a constitutively expressed actin (positive control). The results showed that plants displaying the PRDT1 phenotype over-expressed the mRNA for the PRDT1 gene, indicating the enhanced expression of the PRDT1 gene is correlated with the PRDT1 phenotype.

EXAMPLE 6

Recapitulation of Pathogen Resistance Phenotype

- [0078] The coding sequences of the PRDT1 gene were cloned behind the strong constitutive CsVMV promoter and transformed into wild-type Col-0 Arabidopsis thaliana plants.

 Primary transformants (T1 plants) were selected and allowed to go to seed. Seed from these plants were planted and evaluated for resistance to Peronospora.
- [0079] Approximately 54 individual seedlings from twenty different primary transformants were planted, stratified and grown for 1 week in a growth chamber. The 1 week old seedlings were inoculated with Peronospora as described previously. After 1 week the seedlings were scored for Peronospora infection by counting the number of conidophores on a cotyledon. Plants with no conidiophores on a cotyledon were scored 0, plants with 1-5 were scored 1, plants with 6-10 were scored 2, plants with 10-20 were scored 3 and plants with more than 20 conidophores on a cotyledon were scored 4. Wild-type Col-0 seedlings were included as a control. Six lines over-expressing PRDT1 were significantly more resistant to Peronospora than wild-type Col-0 plants. In the three most resistant lines, over 70% of the plants scored 0 or 1 while in wild-type Col-0 plants only 4% of the plants scored 0 or 1. These results indicate that over-expression of PRDT1 in wild-type Col-0 confers resistance to *Peronospora parasitica*.

Confirmation of Drought Tolerance Phenotype/Genotype Association

[0080] RT-PCR analysis showed that the PRDT1 gene was specifically overexpressed in plants from the line displaying the improved drought tolerance phenotype. Specifically, RNA was extracted from tissues derived from plants exhibiting the PRDT1 phenotype and from wild type COL-0 plants. RT-PCR was performed using primers specific to the sequence presented as SEQ ID NO:1 and to other predicted genes in the vicinity of the T-DNA insertion. The results showed that plants displaying the PRDT1 phenotype over-expressed the mRNA for the PRDT1 gene by about 500 fold compared to wild type plants, indicating that the enhanced expression of the PRDT1 gene is correlated with the PRDT1 phenotype.

EXAMPLE 8

Recapitulation of the Drought Tolerant Phenotype

- [0081] The coding sequences of the PRDT1 gene were cloned behind the strong constitutive CsVMV promoter and transformed into wild-type Col-0 Arabidopsis thaliana plants.

 Primary transformants (T1 plants) were selected and allowed to go to seed. Seed from these plants were planted and evaluated for drought tolerance.
- [0082] Approximately 18 seed from 2 T2 populations (PRDT1-1 and PRDT1-2) transformed with the CsVMV::PRDT1 transgene were planted in soil in 4 inch pots. The seed were stratified at 4° C for 3 days and then grown in the growth room for 4 weeks. Drought stress was imposed on the plants by withholding water 28 days after transfer to the growth room, at this time the plants were beginning to bolt.
- [0083] Observations on every individual plant in each pot were made after the last watering (day 0), seven days after the last watering (day 7), fourteen days after the last watering (day 14) and seventeen days after the last watering (day 17). Plants that were fully turgid were scored 0, plants that showed slight wilting were scored 1, plants that showed significant wilting were scored 2 and plants that were fully desiccated were scored 3.
- [0084] PRDT1-1 and PRDT1-2 plants were significantly more tolerant to drought than wild-type Col-0 plants. After seventeen days without water, Col-0 plants had an average score of 2.7 indicating that most plants were either significantly wilted or fully desiccated. However, the average score of PRDT1-1 and PRDT1-2 plants was 0.1 and 0.4 respectively. These

scores indicate that most of the plants were fully turgid or slightly wilted. These results indicate that overexpression of PRDT1 confers a drought tolerant phenotype.